A strong law for the oscillation modulus of the multivariate empirical process

Einmahl, J.H.J.; Ruymgaart, F.H.

Published in:
Statistics and Decisions

Publication date:
1985

Statistics & Decisions 3, 357–362 (1985)

A STRONG LAW FOR THE OSCILLATION MODULUS OF THE MULTIVARIATE EMPIRICAL PROCESS

J.H.J. Einmahl and F.H. Ruymgaart

Received: Revised version: October 17, 1984

Abstract. A theorem of Mason, Shorack & Wellner [5] on the almost sure behavior of the oscillation modulus of the uniform empirical process is generalized to the multivariate case.

1. INTRODUCTION AND NOTATION

Let $X_1, X_2, \ldots$ be i.i.d. random vectors in $[0,1]^d$, $d \in \mathbb{N}$, that are uniformly distributed over $[0,1]^d$. The empirical process based on the first n vectors is written as

$$(1.1) \qquad U_n = \{U_n(t) = n^{\frac{1}{2}}(\hat{F}_n(t)-|t|),\ t \in [0,1]^d\},$$

where $\hat{F}_n(t)$ is the empirical d.f. defined in the usual way by $n\hat{F}_n(t) = \#\{1 \le i \le n : X_i \in [0,t_1]\times\ldots\times[0,t_d]\}$ and where $|t| = t_1 \times \ldots \times t_d$, with t_j the j-th component of t. The half-open rectangles $(s_1,t_1]\times\ldots\times(s_d,t_d]$ will be written as $R(s,t) = R$ and the class of all such half-open rectangles contained in the unit square is denoted by

$$(1.2) \qquad \mathcal{R} = \{R(s,t) : R(s,t) \subset [0,1]^d\}.$$

AMS 1980 subject classifications: Primary 60F15; Secondary 60G17.
Key words and phrases: multivariate empirical process, oscillation modulus.

Given any (random) function $\Lambda : [0,1]^d \to \mathbb{R}$ and an arbitrary rectangle $R(s,t) \in \mathcal{R}$ we write

$$(1.3) \qquad \Lambda\{R(s,t)\} = \Delta_s^t \Lambda,$$

where Δ_s^t is the usual difference operator, and the Lebesgue measure of the rectangle is written as

$$(1.4) \qquad |R(s,t)| = (t_1 - s_1) \times \ldots \times (t_d - s_d).$$

There are various ways to extend the notion of modulus of continuity or oscillation modulus to higher dimensions. One possible definition that plays a role in the tightness of the empirical processes is

$$(1.5) \qquad \sup_{\|s-t\| \le h} |U_n(s) - U_n(t)|, \quad h \in (0,\infty).$$

For applications to multivariate density estimation, however, the definition

$$(1.6) \qquad \omega_n(h) = \sup_{R \in \mathcal{R}_h} |U_n\{R\}|,$$

where

$$(1.7) \qquad \mathcal{R}_h = \{R(s,t) \in \mathcal{R} : \max_{j=1,\ldots,d} (t_j - s_j) \le h\}, \; h \in (0,\infty),$$

is appropriate. In this note we will exclusively deal with (1.6) and determine its almost sure behavior in the case where $h = h_n$ depends on the sample size.

Attention will be further restricted to the study of $\omega_n(h)$ for sequences $\{h_n\}$ satisfying

$$(1.8) \qquad h_n = (c\, n^{-1} \log n)^{1/d}, \text{ for some } c \in (0,\infty);$$

note that $R \in \mathcal{R}_{h_n}$ entails $|R| \le cn^{-1} \log n$ for such h_n. The restriction is not imposed by the limitations of our approach, but because as far as we know for these sequences the oscillation modulus has not yet been considered in the literature; cf. e.g. Alexander [1] or Stute [7]. In the one dimensional case a similar gap existed that was filled by Mason, Shorack & Wellner [5].

For sequences of the order (1.8) the precise exponential bound for binomial tail probabilities as given in Bennett [2] is indispensable; cf. Stute [7], p. 364. This exponential bound is an essential tool for our basic fluctation inequality that will be given now to conclude this section. We do not give a proof of this inequality but just mention that it is a stronger version of Theorem 1.1 in Ruymgaart & Wellner [6]. For some details of the proof see Einmahl [3]. Let $\psi : [0,\infty) \to [0,\infty)$ be the decreasing function defined by

$$(1.9) \qquad \psi(\lambda) = 2\lambda^{-2} \int_0^\lambda \log(1+x)dx, \ \lambda > 0; \ \psi(0) = 1.$$

Note that $\psi(\lambda) \downarrow 0$ as $\lambda \uparrow \infty$. Let F be *any* d.f. concentrating mass 1 to $[0,1]^d$.

THEOREM 1.1. *Let* $R \in \mathcal{R}$ *with* $F\{R\} \le \frac{1}{2}$. *For each* $\varepsilon \in (0,1)$ *there exists* $C = C(d,\varepsilon) \in (0,\infty)$ *such that*

$$(1.10) \qquad P(\sup_{S\subset R} |U_n\{S\}| \ge \lambda) \le C \exp\left(\frac{-(1-\varepsilon)\lambda^2}{2F\{R\}} \psi\left(\frac{\lambda}{n^{\frac{1}{2}}F\{R\}}\right)\right), \quad \lambda \ge 0,$$

where the supremum is taken over all $S \in \mathcal{R}$ *with* $S \subset R$.

2. THE STRONG LIMIT

Once the basic fluctuation inequality (1.10) is obtained only minor modifications of Mason, Shorack & Wellner [5] are required. As in that paper we will first show how the global supremum in (1.6) is related to a maximum over an only finite number of local suprema. For a vector $x = (x_1,\dots,x_d) \in \mathbb{R}^d$ let us briefly write $\langle x_j\rangle$ if the coordinates should be displayed. Let us note that, for arbitrary $m \in \mathbb{N}$, each $R(s,t) \in \mathcal{R}_h$ with

$$(2.1) \qquad [ms_j] = k(j) \in \{0,1,\dots,m-1\},$$

satisfies

$$(2.2) \qquad R(s,t) \subset R\left(\left\langle\frac{k(j)}{m}\right\rangle, \left\langle\frac{k(j)+1}{m} + h\right\rangle\right);$$

[x] denotes the largest integer $\le x$. Hence it follows that

$$(2.3) \qquad \omega_n(h) \le \max \sup_{R\subset R\left(<\frac{k(j)}{m}>,<\frac{k(j)+1}{m}+h>\right)} |U_n\{R\}| + n^{\frac{1}{2}}dm^{-1}(h+m^{-1})^{d-1},$$

where the maximum is taken over all vectors $(k(1),\ldots,k(d)) \in \{0,1,\ldots,m-1\}^d$.

Let $\varepsilon \in (0,1)$ be the same number as in (1.10) and let us choose $m \in \mathbb{N}$ such that

$$(2.4) \qquad \tfrac{1}{2}\varepsilon^2 h \le m^{-1} \le \varepsilon^2 h.$$

This entails that

$$(2.5) \qquad \tfrac{1}{2}n^{\frac{1}{2}}d\varepsilon^2(1+\tfrac{1}{2}\varepsilon^2)^{d-1}h^d \le n^{\frac{1}{2}}dm^{-1}(h+m^{-1})^{d-1} \le n^{\frac{1}{2}}d\varepsilon^2(1+\varepsilon^2)^{d-1}h^d = C'(d,\varepsilon)n^{\frac{1}{2}}h^d,$$

$$(2.6) \qquad h^d \le \left|R\left(<\frac{k(j)}{m}>,<\frac{k(j)+1}{m}+h>\right)\right| \le h^d(1+\varepsilon^2)^d.$$

It follows from (2.3)-(2.6) that

$$(2.7) \qquad P(\omega_n(h) \ge \lambda h^{\frac{1}{2}d}) \le \Sigma P\left(\sup_{R\subset R\left(<\frac{k(j)}{m}>,<\frac{k(j)+1}{m}+h>\right)} |U_n\{R\}| \ge \lambda h^{\frac{1}{2}d} - C'(d,\varepsilon)n^{\frac{1}{2}}h^d\right),$$

where the summation extends over all $(k(1),\ldots,k(d)) \in \{0,1,\ldots,m-1\}^d$ and where $C'(d,\varepsilon)$ is defined by (2.5).

LEMMA 2.1. *Let* $\varepsilon \in (0,1)$ *be the same number as in* (1.10) *and let* U_n *be the empirical process in* (1.1). *For each* $h \in (0,\frac{1}{4})$, $n \in \mathbb{N}$ *and* $\lambda \ge \varepsilon^{-1}C'(d,\varepsilon)n^{\frac{1}{2}}h^{\frac{1}{2}d}$ *we have*

$$(2.8) \qquad P(\omega_n(h) \ge \lambda h^{\frac{1}{2}d}) \le C(d,\varepsilon)\left(\frac{2}{\varepsilon^2 h}\right)^d \exp\left(-\tfrac{1}{2}\lambda^2(1-\varepsilon)^{3+d}\,\psi\left(\frac{\lambda}{n^{\frac{1}{2}}h^{\frac{1}{2}d}}\right)\right).$$

PROOF. First observe that for λ satisfying the condition it follows that

$$(2.9) \qquad \lambda h^{\frac{1}{2}d} - C'(d,\varepsilon)n^{\frac{1}{2}}h^d \ge \lambda h^{\frac{1}{2}d}(1-\varepsilon).$$

Application of Theorem 1.1 along with (2.6) and the decreasing character of ψ yields the upper bound

(2.10) $$C(d,\varepsilon)\ \exp\left(\frac{-(1-\varepsilon)^3\lambda^2}{2(1+\varepsilon^2)^d}\ \psi\left(\frac{\lambda}{n^{\frac{1}{2}}h^{\frac{1}{2}d}}\right)\right),$$

for each of the probabilities in the summation on the right in (2.7). Finally note that the number of terms in the summation is bounded above by $(2/(\varepsilon^2 h))^d$. Q.E.D.

In the notation of Mason, Shorack & Wellner [5], let β_c^+ be the number, defined for each $c \in (0,\infty)$ by

(2.11) $$\beta_c^+(\log \beta_c^+ - 1) + 1 = c^{-1}, \quad \beta_c^+ > 1.$$

THEOREM 2.1. *Choose* h_n *as in* (1.8) *and let* β_c^+ *be defined for the* c *in* (1.8) *according to* (2.11). *Then we have*

(2.12) $$\lim_{n\to\infty} \frac{\omega_n(h_n)}{(2h_n^d \log(1/h_n^d))^{\frac{1}{2}}} = (c/2)^{\frac{1}{2}}(\beta_c^+ - 1), \text{ a.s.}$$

PROOF. Since Inequality 2 of Mason, Shorack & Wellner [5] carries over immediately to arbitrary dimensions we have, with Lemma 2.1, all the tools to proceed in the same way as in that paper, in order to show that the limsup of the expression on the left in (2.12) is bounded above by the number on the right. The derivation of the fact that the liminf is bounded below by the same number uses heavily (3.11) of Komlós, Major & Tusnády [4]. This derivation is tedious and will be omitted. Q.E.D.

REFERENCES

[1] Alexander, K.S.: Rates of growth and sample moduli for weighted empirical processes indexed by sets. Technical report. University of Washington, Seattle (1984).

[2] Bennett, G.: Probability inequalities for the sum of independent random variables. *J. Amer. Statist. Assoc.* *57*, 33-45 (1962).

[3] Einmahl, J.H.J.: A general form of the law of the iterated logarithm for the weighted multivariate empirical distribution function. Report 8418, Math. Inst., Katholieke Un., Nijmegen (1984).

[4] Komlós, J., Major, P. & Tusnády, G.: Weak convergence and embedding. *Coll. Math. Soc. János Bolyai 11: Limit Theorems of Probability Theory*, 149-165, Amsterdam, North-Holland (1975).

[5] Mason, D.M., Shorack, G.R. & Wellner, J.A.: Strong limit theorems for oscillation moduli of the uniform empirical process. *Z. Wahrscheinlichkeitstheorie Verw. Gebiete 65*, 83-97 (1983).

[6] Ruymgaart, F.H. & Wellner, J.A.: Some properties of weighted multivariate empirical processes. *Statist. Decisions 2*, 199-223 (1984).

[7] Stute, W.: The oscillation behavior of empirical processes: The multivariate case. *Ann. Probability 12*, 361-379 (1984).

J.H.J. Einmahl and F.H. Ruymgaart
Dept. of Mathematics
Catholic University
Toernooiveld
6525 ED Nijmegen
The Netherlands